Exhibit D

High-Frequency Representation of a Single V_H Gene in the Expressed Human B Cell Repertoire

By A. K. Stewart,* C. Huang,‡ B. D. Stollar,‡ and R. S. Schwartz*

*From the *Division of Hematology and Oncology, Department of Medicine, New England Medical Center Hospitals; and the ‡Department of Biochemistry, Tufts University School of Medicine, Boston, Massachusetts 02111*

Summary

Idiotype (Id) 16/6 marks a variable (V) region structure that occurs frequently in the human immunoglobulin repertoire. The basis of the Id has been traced to a germline heavy chain gene segment, V_H18/2 (V_H26). To pursue the molecular basis for the frequency of Id 16/6, we have analyzed polymerase chain reaction–generated Cμ, Cγ, and V_H3 family V gene libraries derived from the circulating and tonsillar B cells of four normal individuals and from the B cells of two patients with active systemic lupus erythematosus (SLE). The frequency of V_H18/2 in these libraries was compared with three control V_H genes, V_H56P1, V_H21/28, and V_HA57. Plaque lifts from Cμ and Cγ V_H cDNA libraries were screened with gene-specific oligonucleotide probes. The frequency of V_H18/2 ranged from 4 to 10% of J_H^+ plaques (two to five times that of control V_H genes). In four V_H3 family-specific libraries derived from rearranged DNA, V_H18/2 represented 19–33% of V_H3^+ plaques. Hybridizing V_H18/2 plaques were 98–100% homologous to the germline V_H gene; mutations when present were often in framework 3. Extensive variation was seen in the complementarity determining region 3 sequences of these rearranged V genes. The high frequency of V_H18/2 expression in the B cell repertoire was confirmed by sequencing randomly picked J_H^+ plaques. In two patients with active SLE the frequency of use of V_H18/2 was not greater than that observed in normal subjects. These results show that V_H18/2 is overrepresented in the B cell repertoire of normal subjects and suggest that the immune repertoire may be dominated by relatively few V genes.

Id 16/6, an idiotypic marker identified originally in a human IgM anti-DNA mAb, has been found in the serum of most patients with active SLE, in the renal and skin lesion of lupus, in the serum of patients with certain bacterial infections or autoimmune diseases other than SLE, in 10% of monoclonal gammopathies, and in normal serum. The Id occurs in a wide spectrum of racial and ethnic groups, and B cells from cord blood, children, and adults can all produce it (reviewed in reference 1). Id 16/6 thus marks a V region structure that occurs in high frequency in the human Ig repertoire. The basis of the Id has been traced to a germline heavy chain gene segment, V_H18/2 (also called V_H26 and V_H30p1), a member of the V_H3 family (2). V_H18/2 has been localized to a 500-kb region in the 3' end of the V_H locus (3). The germline gene has a unique defining sequence in the 5' region of its CDR2, and an oligonucleotide probe corresponding to this region was shown to hybridize to a single 2.0-kb band in genomic DNA (4). This probe has identified V_H18/2 in the genomic DNA of >98% of tested subjects. Moreover, all genomic clones identified by moderate stringency hybridization with the CDR2 probe had an identical sequence (5).

V_H18/2 is thus highly conserved and its Ig product is found in a variety of normal and pathological conditions. To test the possibility that V_H18/2 is overrepresented in B cell populations, we studied the frequency of this gene in the expressed V_H gene repertoire of two normal individuals. To survey the expressed human V_H gene repertoire, we generated Ig cDNA libraries from B cells obtained from normal adults in a two-step PCR procedure without B cell selection or manipulation. V gene primers are not used in this procedure, thus allowing random amplification of all V_H families (6). The frequency of V_H18/2 use was also determined in V_H3 family–specific libraries generated by the PCR of rearranged Ig DNA from the peripheral blood of a normal adult, a tonsil, and from the blood of two patients with SLE. The frequency of V_H18/2 in these libraries was determined by hybridization to plaque lifts with the gene-specific probe and confirmed by sequence analysis. The results with V_H18/2 were compared with two other V_H3 family genes, V_H56p1 and V_HA57, and with V_H21/28 (a member of the V_H1

 J. Exp. Med. © The Rockefeller University Press • 0022-1007/93/02/0409/10 $2.00
Volume 177 February 1993 409–418

family). The results showed a prominence of peripheral blood B cells that had rearranged V_H18/2, and suggest that the V_H gene repertoire of human B cells is strongly biased.

Materials and Methods

Ig cDNA Libraries. PBMC isolated from two normal adult donors (Caucasian and Asian) by centrifugation through Ficoll-Hypaque were washed twice in PBS. No further manipulation was carried out before extraction of mRNA over an oligo(DT) column (Invitrogen, San Diego, CA). Double-stranded (ds)cDNA was synthesized from mRNA according to the method of Gubler and Hoffman (7) and blunt ended with T4 DNA polymerase. The primer for cDNA synthesis was complementary to a sequence within the $C\mu1$ or $C\gamma1$ regions. Two steps of PCR amplification were performed, as described previously (6). The first step was primed by oligonucleotide primers attached to the ends of the ds cDNA. The products were ligated into M13mp19 replicative (RF) DNA. A second amplification used a downstream nested $C\mu$ primer and an upstream primer within the M13 vector DNA. The second PCR products were again ligated to M13RF DNA. This ligation mixture was transformed into DH5α bacteria to form the cDNA library for screening. The M13 plaques were lifted onto Genescreen membranes (DuPont-New England Nuclear, Boston, MA), and the membranes were prehybridized, hybridized, and washed at high stringency as described by Treppichio and Barrett (8). Radiolabeled probes were stripped from the membranes before rehybridization. Plaque lifts were screened by hybridization to a degenerate J_H gene oligonucleotide probe. The J_H probe was end-labeled by T4 polynucleotide kinase and γ-[^{32}P] according to Manniatis et al. (9). Oligonucleotide probes (Fig. 1) complementary to the conserved framework 3 (Fr3)[1] regions of the V_H1 and V_H3 families and to the CDR regions of the individual V_H genes (V_H18/2, V_H21/28, V_H56p1, and V_HA57) were synthesized (Oligos etc. Inc., Wilsonville, OR), and these overlapping oligonucleotides were labeled by filling in of the ends with the Klenow fragment of DNA polymerase 1 and α-[^{32}P]deoxynucleoside triphosphates. Unincorporated nucleotides were removed on NENSORB columns (DuPont-New England Nuclear). For more detailed analysis hybridizing plaques were picked for sequencing by chain termination with dideoxynucleoside triphosphates and sequenase (U.S. Biochemical Co., Cleveland, OH). The resulting sequences were compared with published sequences in the human GenBank database with the FASTA program of the GCG software package (10).

V_H3 Family–specific Libraries. DNA was extracted from lymphoid cells by proteolysis, phenol/chloroform extraction, and precipitated in ethanol. PCR amplification was carried out using a V_H3 leader sequence primer (GCTCTAGAACCATGGAGTTTGGGCTGAG) and a consensus J_H primer (GGGAATTCTGAGGAGACGGTGACCAGGGT). The primers contained XbaI and EcoRI restriction sites to facilitate cloning. The conditions were a 5-min denaturation at 98°C, followed by 94°C for 1 min, 55°C for 1.5 min, and 72°C for 1.5 min for 40 cycles with a 10-min extension at 72°C to finish. The resulting band was cut from low melting–temperature agarose gel. Transformation into DH5αF' *Escherichia coli* was performed, and plaque lifts were screened as described above.

Results

Specificity of Oligonucleotide Probes. The frequency of hybridization of V_H18/2 and control probes was determined in

```
VH3       5'CTCCAGAGACAATTCCAAGAACACGCTGTA 3'
(FR3)                         3'GTTCTTGTGCGACATAAACGTTTACTTGTC 5'

VH1       5' GACCAGGGACACGTCCATCAGCACAGCCTA 3'
(FR3)                          3' TAGTCGTGTCGGATGTACCTCGACTCGTCG 5'

VH18/2    5' GTTATTAGTGGTAGTGGT 3'
(5'cdr2)                  3' TCACCATCACCACCATCG 5'

VHA57     5' GTTATTTATATCATATAT
(5'cdr2)            3' TATAGCACCATCGCGTA 5'

VH56p1    5' GTTATATCATATGATGGA 3'
(5'cdr2)            3' ATATACTACCTTCTTAT 5'

VH21/28   5' ACGCATCCAATGCATAGCATAGCT 3'
(cdr1)                     3'TACGTATCGTATCGATCACTTCGA 5'

VH21/28   5' GAACTTCTGGGAAT 3'
(cdr2)              3' GACACTTATAAAACAT 5'
```

Figure 1. Oligonucleotide probes used to screen libraries. Labeling was performed using a fill-in reaction.

the PCR-generated Ig-specific libraries. Controls included: V_H56p1, a V_H3 family member found initially in fetal liver B cells (11); V_H21/28, a germline V_H1 family member found in autoantibodies (12); and V_HA57, which is most likely a somatically mutated variant of V_H18/2 (4). The specificity of the oligonucleotide probes for the V_H18/2 and V_H21/28 gene segments has been demonstrated previously (4). These experiments showed that an oligonucleotide probe complementary to the 5' end of the CDR2 of V_H18/2 hybridized to a single 2.0-kb band on a Southern blot of digested genomic DNA. Sequence analysis has confirmed that hybridization identifies a single germline V_H gene (4). The present experiments confirmed the specificity of the V_H18/2 probe, which at high stringency identified only plaques with 97–100% homology to V_H18/2 (see below). Hybridization with a combination of probes, identifying both the CDR1 and CDR2 of V_H21/28, had a similar high specificity (D. Rubinstein, unpublished data). The oligonucleotide probes for V_H56p1 and V_HA57 were complementary to their 5' CDR2 regions. Our V_H oligonucleotide overlapped two previously described oligonucleotide probes, M10 (13) and H61 (14). At high stringency both M10 and H61 identify the two bands on TaqI-digested DNA that contain the V_H56p1 germline gene (hv3005) and the highly related GLSJ2 germline gene (13, 14). This area also shares identity with the closely related yet independent germline gene V_H1.9III (13). In fact, in our hands, the V_H56p1 oligonucleotide probe, when hybridized at moderate stringency (10°C below the Tm) to a Southern blot of PstI-digested granulocyte DNA from normal donors, revealed three bands of 35, 15, and 5.2 kb. (not shown). The lack of specificity of the V_H56p1 oligonucleotide is therefore likely to overestimate the frequency of expression of this gene; furthermore, the use of high-stringency washes may underestimate the frequency of expression of V_H18/2 by excluding expressed V_H genes that have undergone somatic mutation in the region identified by the oligonucleotide.

V_H18/2 Is Expressed at High Frequency in Normal Adults. IgM cDNA libraries were generated from the PBL of two normal individuals (Aμ, Tμ) (15). The resultant library plaque lifts were screened with a consensus J_H probe, conserved family-specific V_H3 and V_H1 probes, and, at high stringency,

[1] *Abbreviation used in this paper:* Fr3, framework 3.

Table 1. *Frequency of Oligonucleotide Hybridization to Cμ and Cγ Libraries from Two Individuals*

Name	J_H	V_H3	18/2	56P1	A57	V_H1	21/28
		%	%	%		%	%
Aμ	468	134 (29)	45 (10)	–	–	140 (35)	–
Aμ.B	344	129 (37)	34 (10)	11 (3)	0	142 (41)	8 (2)
Aγ	1,012	631 (61)	45 (4)	9 (1)	0	–	–
Aμ2	432	149 (34)	33 (8)	–	–	–	–
Tμ	284	114 (40)	14 (5)	7 (2)	0	83 (29)	6 (4)

Libraries Aμ and Aμ2 were from the same individual after a 1-yr interval. Aμ.B and Aγ are from the same starting mRNA sample as Aμ.

with the $V_H18/2$-, $V_H21/28$-, V_H56p1-, and V_HA57-specific oligonucleotides. Differences between the two individuals were noted (Table 1); the $V_H18/2$ probe hybridized to 10% of J_H^+ plaques in Aμ but only to 5% of plaques in Tμ (Fig. 2) (33 and 12% of all V_H3^+ plaques, respectively). By comparison, in both subjects only 2% of all J_H^+ plaques hybridized to the V_H56p1 or $V_H21/28$ probes. The V_HA57 probe did not hybridize to any plaques from either individual. The high frequency of expression of $V_H18/2$ in Aμ was confirmed in a duplicate IgM library generated from the same starting mRNA sample (Aμ.B), in which 10% of J_H^+ plaques also hybridized to the $V_H18/2$ probe. Although the predominence of the V_H1 family persisted in this duplicate library and the frequency of $V_H18/2$ remained constant, the frequency of V_H3 hybridizing plaques was higher in library Aμ.B ($p < 0.01$). Sequencing analysis of 54 randomly picked plaques from both Aμ and Aμ.B demonstrated that 26 and 29% of clones in the respective libraries belong to the V_H3 family. This finding strongly suggests that the observed differences in V_H family distribution reflect variations in hybridization conditions rather than a PCR-induced amplification bias. 2 of 23 clones, in which full V_H sequences were obtained, shared 99.7% identity with $V_H18/2$. A third library (Aμ2) was generated from this same individual after an interval of 11 mo. On this occasion, 8% of plaques hybridized to the $V_H18/2$ probe, demonstrating stability of the overrepresentation with time.

Sequences of randomly picked plaques from these libraries demonstrated that each had a distinct CDR3 sequence, ruling out clone duplication secondary to PCR. 12 $V_H18/2$ hybridizing plaques from these Cμ libraries were sequenced at least through CDR1 (three from Aμ, eight from Aμ.B, and one from Tμ) (Fig. 3). All sequences have an open reading frame (Fig. 4). Little mutation was found in the V_H regions. Three sequenced clones from Aμ (AL1.1, AL1.2, AL1.3) had 1, 10, and 3 mutations in Fr3 and one silent mutation in CDR2. In contrast, in the library obtained 1 yr later (Aμ2), only two of eight clones had any mutation and on this occasion all three mutations occurred in the CDRs. The single sequenced clone from Tμ had five mutations, four in Fr3, and one in CDR2.

Although $V_H18/2$ is highly conserved in the germline,

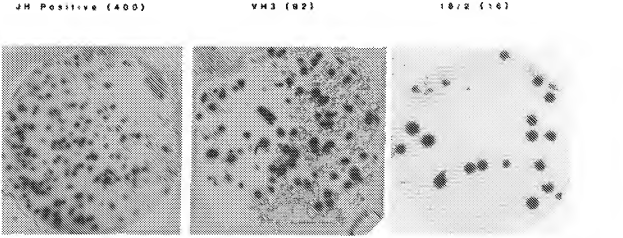


Figure 2. Plaque lifts from Tμ were hybridized sequentially to J_H, V_H3, and $V_H18/2$ probes. On this representative lift V_H3^+ plaques represent 23% of all J_H^+ plaques; $V_H18/2$ hybridized to 17% of V_H3^+ and to 4% of J_H^+ plaques.

Figure 3. Nucleotide sequences of 12 V_H18/2 hybridizing plaques are compared with the germline sequence V_H26. The CDR1 and CDR2 regions are underlined. These sequence data are available from EMBL/GenBank/DDBJ under accession numbers X67060–X67071.

not all members of the V_H3 family have been identified and we were thus unable to conclude definitively that the observed variations from V_H18/2 are indeed due to somatic mutation. Therefore, PCR primers spanning regions of mutation in clones AL1.2 and AL1.3 were used to amplify nonlymphoid DNA. The absence of an amplifiable product supported the probability that the observed base differences were acquired by somatic mutation (not shown). 14 of 23 mutations documented in all 12 sequenced V_H18/2 clones (Aμ, Aμ2, Tμ) resulted in amino acid substitutions, a replacement-

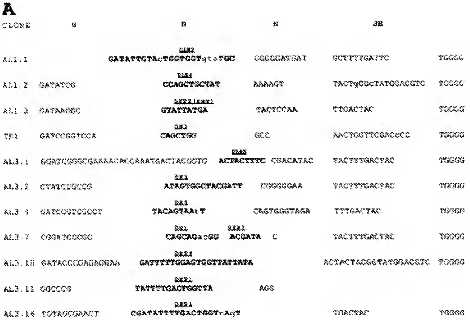
```
                                                 CDR1                             CDR2

18/2     SVQLLESGGGLVQPGGSLRLSCAASGFTFS   SYAMS   WVRQAPGKGLEWVS   AISGSGGSTYYADSVKG
AL1          --------------------------   -----   --------------   -----------------
AL2          --------------------------   -----   --------------   -----------------
AL3      ------------------------------   -----   --------------   -----------------
TK1        ----------------------------   -----   --------------   -----------------
AL3.1                  ----------------   -----   --------------   -----------------
AL3.2                   ---------------   N----   --------------   -----------------
AL3.4                ------------------   -----   --------------   -----------------
AL3.5    ------------------------------   -----   --------------   G----------------
AL3.7          ------------------------   -----   --------------   -----------------
AL3.10                -----------------   -----   --------------   -----------------
AL3.11        -------------------------   -----   --------------   -----------------
AL3.16                  ---------------   -----   --------------   -----------------

                                               CDR3

18/2     RFTISRDNSKNTLYLQMNSLRAEDTAVYYCA
AL1.1    ---------------------------F---   KDIVLVVYAGDDA        FDI  W
AL1.2    --A----D-------------KT-------T   TDIASCYKSYCAM         DV  W
AL1.3    -----------------SA------------   KDKAYYDTPX            DY  W
TK1      ----------S--------------------   RDPVHSWANW           FDP  W
AL3.1    -------------------------------   KDRAKTPNDYGDYFPTY   YFDY  W
AL3.2    -------------------------------   KLSADSGYDSGE        YFDY  W
AL3.4    -------------------------------   KDPSPYSNSVGR         FDY  W
AL3.5    -------------------------------
AL3.7    -------------------------------   RIPPADGRY           YFDY  W
AL3.10   -------------------------------   YDTRSEDFWSGYYNYYGM    DV  W
AL3.11   -------------------------------   KGP                 YFD   W
AL3.16   -------------------------------   KGSELE              YFD   W
```

Figure 4. Predicted amino acid sequences of 12 sequenced V_H18/2 hybridizing plaques.



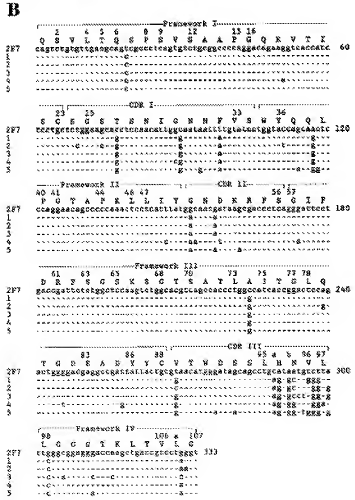


Figure 5. Nucleotide sequences of the CDR3 segments of the 12 V$_H$18/2 hybridizing plaques. Areas with identity to known germline gene segments are shown in bold type, mutations in plain type. These sequence data are available from EMBL/GenBank/DDBJ under accession numbers X67060–X67074.

Table 2. *Frequency of Oligonucleotide Hybridization to V_H3 DNA Libraries from a Normal Adult (AKS), a Tonsil (VAR), and from Two Patients with Active SLE and High-Titer Anti-DNA Antibodies*

Name	V_H3	$V_H18/2$	V_H56P1
		%	%
AKS	177	34 (19)	27 (15)
VAR	79	22 (28)	12 (15)
JAV	80	16 (20)	5 (6)
LG	140	0	-

to-silent ratio of 1.3. Although PCR error may also introduce mutation, our previous work indicated that no more than 1 base in 300 is likely to be misincorporated (15).

In 7 of 11 sequenced $V_H18/2$ genes, 100% identity to known D gene segments could be demonstrated over 9–22 bases in length (Fig. 5). The D gene families DXP (five clones) and DLR (three clones) accounted for 8 of the 11 sequenced clones, a finding that reflects the overrepresentation of these gene families in normal subjects (15–17). The J_H4 gene is also overrepresented in the normal repertoire (15–17) and was used by 7 of the 11 $V_H18/2$ clones described here.

$V_H18/2$ Is Found at High Frequency in V_H3-specific Libraries. In addition to the three $C\mu$ libraries just discussed, V_H3-specific libraries were generated from the lymphocyte DNA of one adult (AKS) and from a tonsillectomy sample (VAR). As seen in Table 2, 19 and 28% of all V_H3^+ plaques in the two libraries hybridized to the $V_H18/2$ CDR2 probe. The V_H56P1 probe hybridized to 15% of the V_H3 plaques in each library. Randomly sequenced plaques demonstrated that these libraries consisted of independent clones. 12 plaques picked at random from AKS were sequenced through CDR1 (Table 3). 3 of the 12 shared 97.4, 99, and 99.6% sequence identity with $V_H18/2$. Three (ks.5, ks.6, ks.17) shared >98% homology with a V_H3 family member (N54P3) found in cord blood (18), two clones (ks.9, ks.19) shared >97% with the fetal cDNA clone M26, and 1 of the 12 (ks.16) had 98% identity with V_H56p1. 10 of 12 rearranged V_H genes found in adult B cells appear to derive from the so-called "fetal repertoire" of V_H genes.

$V_H18/2$ in SLE. The $V_H18/2$ heavy chain is the major determinant of Id 16/6 (2). Levels of this Id fluctuate with disease activity in SLE and can be detected in the characteristic skin and renal lesions of this disease (19, 20). We therefore predicted that, at least in some patients with active disease, $V_H18/2$ would be found at a frequency higher than that observed in normal adults. To test this hypothesis, DNA extracted from the PBLs of two SLE patients with high-titer anti-DNA antibodies (LG and JAV) was PCR amplified with V_H3-specific primers, and packaged in M13 as described above. Hybridization analysis revealed that in patient JAV 20% of V_H3 clones hybridized to $V_H18/2$, and 6% of V_H3 clones to V_H56P1, numbers consistent with those found in normal individuals (Table 2). In patient LG, however, no $V_H18/2$ was detected on hybridization. Indeed, analysis of this patient's serum failed to detect any Id 16/6. Granulocyte DNA from patient LG was amplified with a V_H3 Fr3 and a V_H3 leader primer. A Southern transfer of the product hybridized at high stringency to the $V_H18/2$ 5' CDR2 probe, a finding that implies that this patient has a germline copy of $V_H18/2$ that does not appear in the peripheral B cell population.

$V_H18/2$ in IgG-specific Libraries. As the pathogenic autoantibodies in SLE are high-titer, high-affinity IgG antibodies,

Table 3. *Results of Sequence Analysis of Randomly Picked V_H3 DNA Clones from a Normal Adult (AKS)*

Clone	Homology	Gene bank	V_H bases sequenced	D_H	J_H
	%				
KS.1	99	V_H26	283	DN4	4
KS.3	97.4	V_H26	228	DN1	1
KS.4	99.6	V_H26	252	DXP3	4
KS.9	97.2	M26	253	DLR3	5
KS.19	97.9	M26	283	-	-
KS.16	99.6	FL2-2	279	DLR4	6
KS.5	97.9	N54P3	274	DXP3	4
KS.6	99.6	N54P3	227	-	6
KS.17	98.9	N54P3	282	-	4
KS.10	94.5	V-GL616	249	-	4
KS.20	89.2	63P1	213	DXP'1	6
KS.4	87.3	V_H26	252	DN1	4

KS.16 is 98.6% homologous to V_H56P1.

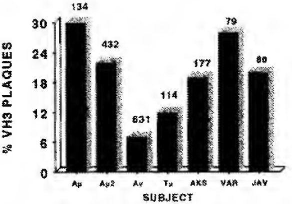


Figure 6. The frequency of $V_H18/2$ as a percentage of V_H3 in the Ig libraries. The number of V_H3^+ plaques studied is shown above each bar. Aμ and Aμ2 libraries were prepared from the same subject, 1 yr apart.

we were interested in the frequency of use and mutation of $V_H18/2$ in the IgG populations of normal donors. An IgG-specific cDNA library was amplified to include all V_H families (Aγ). Hybridization results from this library revealed 45 $V_H18/2^+$ plaques (7% of V_H3 and 4% of all J_H) (Table 1). Four $V_H18/2$ hybridizing plaques were sequenced and were identical in CDR3. As this may invalidate the hybridization data, randomly picked J_H^+ plaques were also sequenced. 3 of 30 randomly picked plaques with independent CDR3 segments were 96, 99, and 100% homologous to $V_H18/2$ at least through CDR1 (not shown). The frequency of $V_H18/2$ expression, as determined by random sequencing, is therefore at least 7% of J_H in this Cγ library.

Discussion

Although 100–200 bands can be seen on Southern blot analysis of human DNA hybridized with V_H family-specific probes, the exact number of functional V_H genes available for rearrangement is unknown (21). The finding of recurrent individual V_H genes in B cell malignancies (22), fetal liver B cells (11, 23), and in autoantibodies produced by adults (24) suggests that some individual V_H gene segments are prone to preferential selection, or alternatively that the number of V_H genes available for use in mature V_HDJ recombinations may be lower than expected (25).

The germline gene $V_H18/2$ is a member of the largest V_H family (V_H3) and is identical to V_H26, which was originally identified on the basis of its homology to a murine V_H gene probe (26). Subsequently, V_H26 has been identified with surprising frequency in fetal liver B cells (10, 21), anti-DNA and other autoantibodies (27, 28), B cell malignancies (29), and in the antibody response of normal individuals to immunization with hemophilus B (30, 31). However, since the representation of $V_H18/2$ in the repertoire of normal B cells was hitherto unknown, interpretation of the overrepresentation of this gene in pathologic conditions was difficult.

The $V_H18/2$ gene segment is demonstrated here to be overrepresented in the unstimulated peripheral blood B cells of normal individuals, with 4–10% of all J_H^+ plaques and 12–33% of all V_H3^+ plaques hybridizing to a $V_H18/2$-specific probe (Fig. 6). Thus, although the V_H3 family is estimated to contain at least 25 members, a single V_H gene accounts for up to 25% of all expressed members of this family. Moreover, this value of 25% is likely to be an underestimate of the true frequency of $V_H18/2$ because the high-stringency conditions used to identify the gene would miss mutant variants in CDR2. The frequency of rearrangement of another recurring V_H3 family member, studied for comparative purposes (V_H56p1), was 2% of J_H^+ and 6–15% of V_H3^+ plaques. As a further comparison, $V_H21/28$, a member of the large V_H1 family, also represented 2% of all J_H^+ plaques (6–7% of all V_H1 plaques). Thus, although $V_H18/2$ predominates, all of the studied V_H genes are rearranged at a higher than expected frequency, assuming a total of at least 100 functional V_H genes.

Examples of other V gene sequences identified at high frequency in normal individuals include the V_H4 family member $V_H4.21$ (31), the V_H1 family member V_H51p1 (32) the Vκ gene *humkv*3.25 (33, 34), the D gene segment Dn1 (15–17), and the J_H gene J_H4 (15–17). V_H51p1 and Vh4.21 have been identified by their respective idiotypic markers in 2 and 3% of tonsillar B cells, and $V_H4.21$ in 10% of bone marrow B cells (31, 32). V_H6 and the recently described V_H7 gene segment have been found at high frequency in fetal (11, 23) and cord blood B cells (18), and in up to 6% of all circulating B cells in one normal adult (15). It is likely that other examples of V_H genes found at high frequency in normal individuals remain to be identified. By extension, it is therefore possible that the expressed repertoire of human V genes will ultimately prove to represent only a fraction of the potentially available germline.

A number of reasons have been proposed for the overrepresentation of certain germline V genes in fetal liver, autoantibodies, B cell malignancy, and now in normal individuals. These relate to chromosomal position (35), the number of gene copies in the germline (36), preferential rearrangement on the grounds of unique recombinase accessibility or recognition sequences, the presence of gene-specific promotor enhancer sequences (37), and preferential selection on the basis of antigen binding or Id specificities (38).

A gene product may be found at higher than predicted frequency if more than one copy exists in the germline. In fact, there are probably two copies of $V_H18/2$ in the germline (5), and this may partially explain its preponderance in the repertoire. However, as our V_H56p1 probe recognizes at least two highly related yet independent germline V_H genes (13, 14), the predominance of $V_H18/2$ over V_H56p1 (and related genes) in this study implicates factors other than multiple copies.

$V_H18/2$ is highly conserved and its coding region is not polymorphic (5). This suggests that $V_H18/2$ may be preferentially selected because of the antigen binding or idiotypic properties of its protein product. $V_H18/2$ has a sequence in Fr3 that is highly conserved between species and within the

V_H3 family, but that differs from other families (37). This sequence may encode a unique antigen binding site in the protein product not related to the classical CDR-related binding sites. Of interest, the consensus sequence from all known V_H3 family members is identical to the $V_H18/2$ sequence in this region. It is therefore notable that 9 of 11 $V_H18/2$ amino acid substitutions found in this study were in Fr3 (Figs. 4 and 5). Perhaps $V_H18/2$ is selected on the basis of this highly conserved potential antigen binding site. Precedence for this hypothesis has been demonstrated in mice in which clonal persistence of B lymphocytes in normal animals is determined by V_H family-dependent selection (39).

Since Id16/6 is abundant in the serum of some patients with active SLE and its levels fluctuate with disease activity, we expected to find an excess of $V_H18/2$ plaques in some patients with active SLE. In a V_H3 family library amplified from the DNA of a patient (JAV) with active lupus, $V_H18/2$ was found to be rearranged in 20% of V_H3^+ clones, a frequency similar to that of normal individuals. In a second patient (LG), a patient without detectable Id in the serum, no $V_H18/2$ was identified by hybridization. By PCR analysis this patient appears to have a copy of $V_H18/2$ in the germline, and the absence of rearranged $V_H18/2$ in this patient remains unexplained.

It therefore appears, at least from these limited studies, that a difference in the frequency of use of $V_H18/2$ cannot explain the elevated levels of Id 16/6 in patients with active SLE. Our observation could be explained if other V_H genes contribute to the Id (2), if activation of normally quiescent B cells bearing an 18/2 rearrangement results in the release of the Id into the serum, or if plasma cells producing Id 16/6 are sequestered from the circulation.

A possibility not previously considered is that the number of germline V_H genes used in rearrangements is far lower than supposed. Such a limitation in diversity has been demonstrated in other species, such as the chicken and rabbit, which both generate diversity using an extremely restricted set of V_H genes (40, 41). The chicken uses only one V_H and one $V\lambda$ gene. Indeed, all other $V\lambda$ genes in the germline of the species are pseudogenes, portions of which are subsequently used in gene conversion events to generate antibody diversity. In the rabbit, only one of many functional V_H genes is used. This gene (V_H1), which is the most 3' V_H gene, generates diversity by a combination of somatic mutation and gene conversion. While gene conversion events are probable in humans their presence has yet to be conclusively documented (16).

The majority of the $C\mu$ $V_H18/2$ clones we sequenced showed little mutation, even in CDR3, suggesting that the population of IgM^+ B cells bearing $V_H18/2$ rearrangements may form part of the naive immune repertoire. Unmutated $V_H18/2$ was also found in a $C\gamma$ library. Taken together with the demonstration that $V_H18/2$ can encode the heavy chains of anti-DNA antibodies, this finding suggests that germline genes capable of forming naturally occurring autoantibodies are not deleted from the IgG population of B cells in normal individuals.

Our results suggest that a significant fraction of the human Ig repertoire originates from a preimmune repertoire that is dominated by relatively few V genes. In this regard the remarkable polyspecificity of antibodies encoded by germline V genes may be important. A limited number of polyreactive clones could form the substrate from which a diverse repertoire arises after clonal selection (24).

The recurrent presence of individual germline V_H genes can now be extended from the restricted B cell populations of the fetal repertoire, autoantibodies, and B cell malignancies to the expressed V gene repertoire of normal adults. We postulate that only a fraction of available germline V genes are used recurrently in the expressed repertoire, and that polyspecificity of naturally occurring antibodies in combination with CDR3 and somatic mutation compensate for the restriction to antibody diversity. The mechanisms by which preferential use of an individual V_H gene arise remain speculative and deserve further investigation.

We thank D. Rubinstein and K. Barrett for helpful discussions and the use of their oligonucleotide probes, and L. Tucker for performing the Id 16/6 ELISA. The secretarial help of V. Barthel is gratefully acknowledged.

This research was supported by National Institutes of Health grants AI-26450 and AI-28899. A. K. Stewart is a fellow of the Medical Research Council of Canada.

Address correspondence to R. S. Schwartz, Division of Hematology-Oncology, Box 245, N.E.M.C., 750 Washington Street, Boston, MA 02111.

Received for publication 19 May 1992 and in revised form 25 September 1992.

References

1. Schwartz, R.S. 1989. Properties of a system of public idiotypes of lupus autoantibodies. *Int. Rev. Immunol.* 5:153.
2. Young, F., L. Tucker, D. Rubinstein, T. Guillaume, J. Andre-Schwartz, K.J. Barret, R.S. Schwartz, and T. Logtenberg. 1990. Molecular analysis of a germ line-encoded idiotypic marker of pathogenic human lupus autoantibodies. *J. Immunol.* 145:2545.
3. Walter, M.A., H.M. Dosch, and D.C. Cox. 1991. A deletion map of the human immunoglobulin heavy chain variable re-

gion. *J. Exp. Med.* 174:335.
4. Guillaume, T., D.B. Rubinstein, F. Young, L. Tucker, T. Logtenberg, R.S. Schwartz, and K.J. Barrett. 1990. Individual VH genes detected with oligonucleotide probes from the complementarity-determining regions. *J. Immunol.* 145:1934.
5. Rubinstein, D.B., T. Guillaume, M. Symann, A.K. Stewart, and K.J. Barrett. 1992. Restriction fragment length polymorphisms and single germline coding sequence in VH18/2, a duplicated gene encoding autoantibody. *Mol. Immunol.* In press.
6. Huang, C., and D.B. Stollar. 1991. Construction of representative immunoglobulin region cDNA libraries from peripheral blood lymphocytes without in-vitro stimulation. *J. Immunol. Methods.* 141:227.
7. Gubler, U., and B. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene (Amst.).* 25:263.
8. Treppichio, W., and K.J. Barrett. 1984. Eleven MRL-lpr/lpr anti-DNA antibodies are encoded by genes from four VH gene families: a potentially biased usage of VH genes. *J. Immunol.* 138:2323.
9. Manniatis, T., E.F. Fritsch, and J. Sambrook. 1989. Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
10. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research.* 12:387.
11. Schroeder, H.W., J.L. Hillson, and R.M. Perlmutter. 1987. Early restriction of the human antibody repertoire. *Science (Wash. DC).* 238:791.
12. Dersimonian, H., R.S. Schwartz, K.J. Barrett, and B.D. Stollar. 1987. Relationship of human variable region heavy chain germline genes to genes encoding anti-DNA antibodies. *J. Immunol.* 139:2496.
13. Willems van Dijk, K., H.W. Schroeder, R.M. Perlmutter, and E.C.B. Millner. 1989. Heterogeneity in the human immunoglobulin VH locus. *J. Immunol.* 142:2547.
14. Sasso, E.H., K. Willems van Dijk, and E.C.B. Millner. 1990. Prevalence and polymorphism of human VH3 genes. *J. Immunol.* 142:2751.
15. Huang, C., A.K. Stewart, R.S. Schwartz, and B.D. Stollar. 1992. Immunoglobulin heavy chain gene expression in peripheral blood B lymphocytes. *J. Clin. Invest.* 89:1331.
16. Sanz, I. 1991. Multiple mechanisms participate in the generation of diversity of human H chain CDR3 regions. *J. Immunol.* 147:1720.
17. Yamada, M., R. Wasserman, B.A. Reichard, S. Shane, A.J. Caton, and G. Rovera. 1991. Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult human peripheral blood lymphocytes-B. *J. Exp. Med.* 173:395.
18. Mortari, F., J.A. Newton, J.Y. Wang, and H.W. Schroeder. 1992. The human cord blood antibody repertoire. Frequent usage of the VH7 gene family. *Eur. J. Immunol.* 22:241.
19. Isenberg, D., C. Dudeny, F. Wojnaruska, and B.S. Bhogal. 1985. Detection of cross reactive anti-DNA antibody idiotypes on tissue bound immunoglobulins from skin biopsies of lupus patients. *J. Immunol.* 135:261.
20. Isenberg, D., and C. Collins. 1985. Detection of cross reactive anti-DNA idiotypes on renal tissue bound immunoglobulins from lupus patients. *J. Clin. Invest.* 76:287.
21. Berman, J.E., S.J. Mellis, R. Pollock, C.L. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C.R. Cantor, and F.W. Alt. 1988. Content and organization of the human Ig VH locus: definition of three new VH families and linkage to the Ig CH locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:727.
22. Logtenberg, T., M.E.M. Scutte, G. Inghirami, J.E. Berman, F.H.G. Gmelig-Meyling, R.A. Insel, D.M. Knowles, and F.W. Alt. 1989. Immunoglobulin VH gene expression in human B cell lines and tumours: biased VH gene expression in chronic lymphocytic leukemia. *Int. Immunol.* 1:362.
23. Schroeder, H.W., and J.Y. Wang. 1990. Preferential utilization of conserved immunoglobulin heavy chain variable gene segments during human fetal life. *Proc. Natl. Acad. Sci. USA.* 87:6146.
24. Pascual, V., and D.J. Capra. 1991. Human immunoglobulin heavy-chain variable region genes: organization, polymorphism, and expression. *Adv. Immunol.* 49:1.
25. Stewart, A.K., C. Huang, D.B. Stollar, and R.S. Schwartz. 1992. VH gene representation in autoantibodies reflects the normal human B cell repertoire. *Immunol. Rev.* 128:101.
26. Matthysesens, G., and T.H. Rabbitts. 1980. Structure and multiplicity of genes for the human immunoglobulin heavy chain variable region. *Proc. Natl. Acad. Sci. USA.* 77:6561.
27. Spatz, L.A., K.K. Wong, M. Williams, R. Desai, J. Golier, J.E. Berman, F.W. Alt, and N. Latov. 1990. Cloning and sequence analysis of the VH and VL regions of an anti-DNA antibody from a patient with peripheral neuropathy and chronic lymphocytic leukemia. *J. Immunol.* 144:2821.
28. Kaburaki, J., and B.D. Stollar. 1987. Identification of human anti-DNA, anti-Sm and anti-SS-A serum antibodies bearing the cross reactive 16/6 idiotype. *J. Immunol.* 139:385.
29. Shoenfeld, Y., O. Ben-Yehuda, Y. Naparstek, and Y. Wilner. 1986. The detection of a common idiotype of anti-DNA antibodies in the sera of patients with monoclonal gammopathies. *J. Clin. Immunol.* 6:194.
30. Silverman, G.J., and A.H. Lucas. 1991. Variable region diversity in human circulating antibodies specific for the capsular polysaccharide of hemophilus-influenzae type B- preferential usage of two types of VH3 heavy chains. *J. Clin. Invest.* 88:911.
31. Adderson, E.E., P.G. Shackelford, A. Quinn, and L.W. Carroll. 1991. Restricted Ig H chain V- gene usage in the human antibody response to hemophilus-influenzae type b capsular polysaccharide. *J. Immunol.* 147:1667.
32. Pascual, V., and J.D. Capra. 1992. VH4-21, a human VH gene segment overrepresented in the autoimmune repertoire. *Arthritis Rheum.* 35:11.
33. Kipps, T.J., and S.F. Duffy. 1991. Relationship of the CD5 B cell to human tonsillar lymphocytes that express autoantibody-associated cross reactive idiotypes. *J. Clin. Invest.* 87:2087.
34. Kipps, T.J., E. Tomhave, L.F. Pratt, S. Duffy, P.P. Chen, and D.A. Carson. 1989. Developmentally restricted VH gene expressed at high frequency in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA.* 86:5913.
35. Yancopoulos, G.D., S.V. Desiderio, M. Paskind, J.F. Kearney, D. Baltimore, and F.W. Alt. 1984. Preferential utilization of the most JH proximal VH gene segments in pre-B cell lines. *Nature (Lond.).* 311:727.
36. Chen, P.P., and P.M. Yang. 1990. A segment of VH gene locus is duplicated. *Scand. J. Immunol.* 31:593.
37. Chen, P.P., R.W. Soto-Gil, and D.A. Carson. 1990. The early expression of some human autoantibody associated heavy chain variable region genes is controlled by specific regulatory elements. *Scand. J. Immunol.* 31:673.
38. Kirkham, P., F. Mortari, J.A. Newton, and H.W. Schroeder. 1992. Immunoglobulin VH clan and family identity predicts variable domain structure and may influence antigen binding. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:603.

39. Thomas-Vaslin, V., L. Andrade, A. Freitas, and A. Coutinho. 1991. Clonal persistence of lymphocytes-B in normal mice is determined by variable region-dependent selection. *Eur. J. Immunol.* 21:2239.
40. Knight, K.L., and R.S. Becker. 1990. Molecular basis of the allelic inheritance of rabbit immunoglobulin VH allotypes: implications for the generation of antibody diversity. *Cell.* 60:963.
41. Reynaud, C.A., V. Anquez, H. Grimal, and J.C. Weil. 1987. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell.* 48:379.